

# **De cross-sectionele en longitudinale ontwikkeling van de intelligentietestscores in PRIMA**

G. Driessen



DE CROSS-SECTIONELE EN LONGITUDINALE ONTWIKKELING  
VAN DE INTELLIGENTIETESTSCORES IN PRIMA



# De cross-sectionele en longitudinale ontwikkeling van de intelligentietestscores in PRIMA

G. Driessen

ITS – Nijmegen

Uitgave en verspreiding:

ITS

Postbus 9048

6500 KJ Nijmegen

Tel.: 024-3653500

<http://www.its.ru.nl>

© 2005 ITS, Radboud Universiteit Nijmegen

Behoudens de in of krachtens de Auteurswet van 1912 gestelde uitzonderingen mag niets uit deze uitgave worden verveelvuldigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze dan ook, en evenmin in een retrieval systeem worden opgeslagen, zonder de voorafgaande schriftelijke toestemming van het ITS van de Radboud Universiteit Nijmegen.

No part of this book/publication may be reproduced in any form, by print, photoprint, microfilm or any other means without written permission from the publisher

# Inhoud

<b>Management samenvatting</b>	1
<b>1 Inleiding</b>	3
1.1 Binnen- en buitenschools	3
1.2 Probleemstelling en onderzoeksvragen	4
<b>2 Steekproeven, variabelen en methode</b>	7
2.1 Steekproeven	7
2.2 Variabelen	8
2.3 Methode	10
<b>3 Resultaten</b>	11
3.1 Ontwikkelingen in intelligentie?	11
3.2 Intelligentie en achtergrondkenmerken	14
3.3 Intelligentie, prestaties en achtergrondkenmerken	14
3.4 De longitudinale ontwikkeling van intelligentie	15
3.5 De voorspelling van de prestaties	16
<b>4 Conclusies</b>	23
<b>Referenties</b>	25
<b>Bijlage</b> – Overwegingen met betrekking tot de opname van een intelligentietest in het vervolg op PRIMA	29





## Management samenvatting

### De cross-sectionele en longitudinale ontwikkeling van de intelligentietestscores in PRIMA

NWO/PROO-projectnummer 411-20-039

In het cohortonderzoek Primair Onderwijs (PRIMA), dat sinds 1994/95 tweejaarlijks wordt uitgevoerd, zijn in de groepen 4, 6 en 8 twee nonverbale intelligentietests afgenomen. In analyses zijn de resultaten op deze tests tot nu toe onderbenut; ze zijn vooral als controlevariabele, maar ook als effectvariabele opgenomen. Er bestaat nog veel onduidelijkheid over hoe de intelligentiescores zich ontwikkelen in de tijd (zowel transversaal als longitudinaal) en wat daarbij de relatie is met achtergrondkenmerken als sekse, sociaal milieu en etniciteit en met taal-, reken- en leesprestaties. Inzicht daarin is enerzijds van belang in verband met secundaire analyses op al beschikbare data, en anderzijds in verband met een nieuw op te starten cohortonderzoek. In dit rapport worden beide aspecten aan de orde gesteld. Dat gebeurt aan de hand van de volgende clusters van vragen:

- Hoe hebben de intelligentietestscores zich over de vijf PRIMA-metingen heen ontwikkeld? Zijn er daarbij verschillen tussen de drie jaargroepen en tussen de twee onderscheiden subtests? Welke samenhang is er met de achtergrondkenmerken van leerlingen?
- Wat is bij de verklaring van verschillen in intelligentiescores de relatieve bijdrage van elk van de achtergrondkenmerken (intelligentie als effectvariabele)? Welke samenhang is er met de cognitieve prestatie-maten?
- Wat voegt opname van intelligentietests in PRIMA toe aan de afname van prestatietoetsen? Wat is de waarde van intelligentie als controlevariabele bij de verklaring van prestatieverschillen?

Voor de analyses is gebruik gemaakt van de gegevens van circa 30000 leerlingen per PRIMA-meting. Qua technieken zijn met name variantie- en regressie-analyses gehanteerd.

Met betrekking tot de vraag naar *ontwikkelingen in intelligentiescores en de relatie met achtergrondkenmerken en toetsprestaties* kan het volgende worden geconcludeerd.

Over de vijf PRIMA-metingen heen zijn er geen verschillen in de hoogte van de scores op de intelligentietests, en dat geldt tevens voor de subtests Figuren en Exclusie, voor de jaargroepen 4, 6 en 8, en voor jongens en meisjes, de onderscheiden ouderlijke opleidingsniveaus en etnische groepen.

Bij de verklaring van verschillen in intelligentiescores – waarbij intelligentie dus als een effectvariabele wordt opgevat – weegt het ouderlijke opleidingsniveau het zwaarste mee, de etnische herkomst is minder relevant, en sekse is helemaal niet van belang.

De intelligentiescores hangen het sterkste samen met de rekenvaardigheid, wat gezien het non-verbale karakter van de tests niet verwondert. De samenhang met begrijpend lezen is minder

sterk en met de taalvaardigheid nog weer zwakker. Deze trend doet zich voor binnen elk van de onderscheiden etnische groepen.

De samenhangen tussen de intelligentiescores in lagere en hogere jaren zijn voor de subtest Figuren matig sterk tot sterk, voor Exclusie zijn ze echter beduidend lager. Dit geldt voor elk van de categorieën van de achtergrondkenmerken. De stabiliteit van de scores over de jaargroepen heen is daarmee vrij laag; ze dienen waarschijnlijk vooral te worden opgevat als een momentopname.

In de voorspelling van de taal-, reken- en leesprestaties in groep 8 heeft intelligentie een substantiële, maar per prestatiedomein verschillende bijdrage; voor rekenen is die het grootst. Echter, als controlevariabele is de betekenis van intelligentie minimaal. Extra controle voor intelligentie, bovenop die van achtergrondkenmerken en eerdere prestaties leidt slechts tot een zeer geringe verbetering van de verklaring van prestatieverschillen.

Met betrekking tot de vraag naar de eventuele *opname van een intelligentietest in het toekomstige cohortonderzoek* kan het volgende worden geconcludeerd.

Gegeven de ervaringen met de tests in PRIMA en VOCL, de zeer beperkte meerwaarde, zowel in termen van controle- als effectvariabele, het niet-beschikbaar zijn van geschikte alternatieven en de onduidelijkheid wat betreft de functie is het de vraag of er überhaupt een intelligentietest zou moeten worden meegenomen. Als er al een zou moeten worden afgenomen, dan zou dat waarschijnlijk een test moeten zijn die ‘fluid intelligence’ meet, die verschillende domeinen bestrijkt en die in de beginfase van het primair onderwijs kan worden afgenomen. Aangezien een dergelijke test voor gebruik in het beoogde cohortonderzoek niet beschikbaar is, is het misschien beter helemaal geen intelligentietest mee te nemen.

# 1 Inleiding

## 1.1 Binnen- en buitenschools

Sinds de start van PRIMA in 1994 is gebruik gemaakt van een intelligentietest. Intelligentie wordt hier opgevat als een latente, niet-rechtstreeks observeerbare eigenschap die bepaalt hoe goed en snel iemand allerlei meer of minder abstracte denkbeelden kan oplossen. Sommigen zien intelligentie als een biologisch of genetisch bepaald gegeven dat op zich vastligt. Anderen zijn van mening dat intelligentie eigenlijk bestaat uit algemene oplossingsstrategieën die door leren en oefening verbeterd kunnen worden. De discussie over de relatieve invloed van het aangeboren potentieel dan wel de omgeving (incl. het onderwijs) en de link naar milieu, ras en cultuur is al oud en laait bij tijd en wijle op (bv. Jensen, 1969; Herrnstein & Murray, 1994; Friesen & Voestermans, 1995). Het is hier niet de plaats om op die ‘nature vs. nurture’ controverse in te gaan; voor een overzicht wordt verwezen naar Neisser e.a. (1990).

Van belang is wel de reden waarom bij het ontwerp van PRIMA en zijn voorganger LEO (Landelijke Evaluatie Onderwijsvoorrrangsbeleid) een intelligentietest is opgenomen. In de praktijk bevatten intelligentietests doorgaans reeksen denkopgaven, vaak met een tijdslimiet. Enerzijds zijn de vaardigheden die vereist zijn om deze opgaven te maken dezelfde vaardigheden als die welke ook nodig zijn bij schoolvakken als taal en rekenen, maar die niet expliciet op school worden onderwezen. De beheersing van die vaardigheden wordt daarmee beïnvloed door wat een kind in de buitenschoolse omgeving krijgt aangeboden of oppikt. Anderzijds is het ook zo dat door de impliciete instructie en oefening tijdens de schoollessen vaardigheden worden geleerd die via een intelligentietest worden gemeten. Er is daarmee een sterke samenhang tussen de resultaten op intelligentietests en schoolprestatietoetsen. Deze samenhang is echter niet perfect: intelligentietests meten niet precies hetzelfde als prestatietoetsen. De intelligentietests kunnen worden opgevat als indicatoren van algemene niet-schoolse cognitieve vaardigheden waarin leerlingen op school ook een zekere oefening krijgen. Het uitgangspunt is daarmee dat de vaardigheden die met de intelligentietests worden gemeten een grotere buitenschoolse component bevatten dan die met de prestatietoetsen worden bepaald. Verondersteld wordt dat door ook intelligentietests op te nemen in het PRIMA-onderzoek het effect van de school, c.q. het onderwijs zuiverder kan worden geschat.

Uit het bovenstaande volgt dat intelligentietestresultaten kunnen worden opgevat als een *controlevariabele* voor vaardigheidsverschillen tussen leerlingen ten gevolge van buitenschoolse invloeden. Om van een zuiver schooleffect te mogen spreken dienen er prestatieverschillen te resteren nadat gecontroleerd is voor de invloed van intelligentie. Het is echter ook mogelijk intelligentie als een *effectvariabele* op te vatten. Intelligentietests bevatten namelijk vaak een talige component, niet alleen als expliciet te meten onderdeel van de test, maar ook als medium voor testinstructie. En ook bevatten ze vaak met rekenen en wiskunde verbonden eigenschappen als ruimtelijk inzicht en ordening. Zowel taal en rekenen (wellicht het laatste nog meer dan het eerste) betreffen vaardigheden die ook op school geleerd worden. De score op een intelligentietest kan dus evenals die op een taal- of rekentoets als een effect van het onderwijs worden be-

schouwd (Tesser, Mulder & Van der Werf, 1991; voor een overzicht van de discussie met betrekking tot causaliteit ofwel de oorzaak-gevolgdiscussie, zie Ceci, 1991).

Bij PRIMA zijn twee nonverbale intelligentietests afgenomen, te weten Figuren samenstellen en Exclusie. Van elke test bestaan drie verschillende versies, namelijk voor groep 4, 6 en 8. Tot nu toe is in PRIMA op twee manieren gebruik gemaakt van de intelligentiegegevens. Allereerst is in de vijf basisrapportages standaard in beschrijvende zin een uitsplitsing gemaakt naar de sociaal-etnische achtergrond en sekse van de leerlingen, waarbij de totaalscore (beide tests samen) is gepresenteerd (bv. Driessen, Van Langen & Vierke, 2004). Hierbij gaat het om het gebruik van intelligentie als een effectvariabele. Daarnaast zijn in een aantal multivariate analyses de test-scores als controlevariabelen meegenomen, doorgaans ook in de vorm van één totaalscore (bv. Driessen, 2002; Luyten & Bosker, 2004; Mulder, 1994; Mulder, Tesser & Vierke, 1994; Mulder, Vierke & Petersen, 1997).

## **1.2 Probleemstelling en onderzoeksvragen**

De resultaten op de intelligentietests zijn in PRIMA onderbenut; het accent ligt steeds op de prestaties van de leerlingen. Onduidelijk is bijvoorbeeld of er sprake is van een ontwikkeling qua intelligentie over cohorten heen en of er daarbij eventueel verschillen zijn naar jaargroep en subgroepen van leerlingen op basis van bijvoorbeeld sekse, opleiding ouders en etniciteit. Mulder (1993) heeft er binnen LEO op gewezen dat er sprake is van een sprong, i.c. inhaalslag van de 1.90-leerlingen tussen groep 4 en groep 6. De indruk bestaat daarmee dat er tussen sommige LEO-metingen wel verschillen optreden, maar hoe groot die precies zijn en waar dat mee te maken heeft is vooralsnog onbekend. Binnen PRIMA zijn deze relaties in het geheel nog niet onderzocht. Eveneens is onduidelijk wat de relatie van de intelligentiescores met de cognitieve prestatiematen, te weten de taal-, reken- en leesvaardigheid, en of er daarbij wellicht verschillen zijn tussen de onderscheiden subtests.

In dit rapport worden deze relaties nader onderzocht, enerzijds met het oog op toekomstige analyses op de reeds beschikbare PRIMA-gegevens, en anderzijds met het oog op de toekomst van het (vervolg) op het PRIMA-cohort:

- In de al uitgevoerde PRIMA-metingen zijn steeds de intelligentietests afgenomen. Dat de betreffende gegevens tot nu toe zijn onderbenut heeft mede te maken met de onduidelijke status van de intelligentietests. Door hier nu meer systematisch onderzoek naar te doen, zou er meer zicht op zijn verkregen. Dat is van groot belang voor toekomstige gebruikers van deze al beschikbare PRIMA-bestanden.
- De PROO wil een nieuw cohortonderzoek opstarten. Een belangrijke vraag daarbij is of er ook intelligentietests moeten worden meegenomen. Meer kennis over de relaties tussen de intelligentiescores, toetsprestaties en achtergrondkenmerken is in verband met de beslissing over handhaving en/of bijstelling van de intelligentietest uiterst relevant.

De centrale variabelen in dergelijke analyses zijn de scores op beide intelligentietests, de taal-, reken- en leesscores, en de leerlingachtergrondkenmerken sekse, opleiding ouders en etniciteit.

Concreet kunnen de volgende clusters van vragen worden onderscheiden:

- Hoe hebben de intelligentietestscores zich over de vijf PRIMA-metingen heen ontwikkeld? Zijn er daarbij verschillen tussen de drie jaargroepen en tussen de twee onderscheiden subtests?<sup>1</sup> Welke samenhang is er met de achtergrondkenmerken van leerlingen?
- Wat is bij de verklaring van verschillen in intelligentiescores de relatieve bijdrage van elk van de achtergrondkenmerken (intelligentie als effectvariabele)? Welke samenhang is er met de cognitieve prestatiematen?
- Wat voegt opname van intelligentietests in PRIMA toe aan de afname van prestatietoetsen? Wat is de waarde van intelligentie als controlevariabele bij de verklaring van prestatieverschillen?

Hierna wordt allereerst verslag gedaan van de analyses die zijn uitgevoerd om bovenstaande vragen te beantwoorden. Dit betreft de vragen zoals ze zijn geformuleerd in de subsidieaanvraag. Daarna wordt ingegaan op de consequenties van de bevindingen voor het al dan niet opnemen van een intelligentietest in het vervolg van PRIMA; dat gebeurt in de vorm van een advies. Omdat dit feitelijk geen deel uitmaakt van de onderzoeksopdracht, is dit advies in een bijlage opgenomen.

---

<sup>1</sup> In het geval er verschillen in intelligentiescores bestaan tussen de vijf metingen, is een bijkomende vraag hoe dergelijke verschillen verklaard kunnen worden.



## 2 Steekproeven, variabelen en methode

### 2.1 Steekproeven

Gebruik is gemaakt van de gegevens van de eerste vijf PRIMA-metingen (P1-P5). PRIMA is een landelijk cohortonderzoek onder circa 600 basisscholen en 60000 leerlingen. Sinds 1994 worden tweejaarlijks gegevens verzameld bij schooldirecteuren, groepsleerkrachten, leerlingen en hun ouders. Voor een technische verantwoording en details wordt verwezen naar Driessen, Van Langen & Vierke (2000, 2002, 2004); Driessen, Van Langen, Portengen & Vierke (1998); Van Langen, Vierke & Robijns (1996); Jungbluth, Van Langen, Peetsma & Vierke (1996).

In PRIMA wordt een representatieve steekproef van circa 420 scholen onderscheiden en een aanvullende steekproef van circa 180 scholen. De eerste is representatief voor alle basisscholen, de tweede bevat scholen met veel allochtone en autochtone achterstandsleerlingen. De aanvullende steekproef is opgenomen om ook betrouwbare uitspraken te kunnen doen over (m.n. allochtone achterstands)groepen die qua aantallen minder goed vertegenwoordigd zijn in de representatieve steekproef. De totale steekproef van 600 scholen kent dus een oververtegenwoordiging van achterstandsscholen. In het onderhavige onderzoek worden twee soorten analyses uitgevoerd, namelijk enerzijds cross-sectionele en transversale ('quasi longitudinaal') en anderzijds longitudinale.

De *cross-sectionele analyses* zijn gericht op het in beeld brengen van de ontwikkelingen qua intelligentie over de cohorten, i.c. PRIMA-metingen en het beschrijven van samenhangen tussen de achtergrondkenmerken, intelligentie en prestaties. Om daar een goed beeld van te geven worden deze analyses verricht op de representatieve steekproef. De aantallen leerlingen variëren per jaargroep van 9000 tot 10000 per meting.

Met behulp van *longitudinale analyses* worden leerlingen gedurende meerdere meetmomenten gevolgd. Het gaat dan bijvoorbeeld om de relatie tussen de intelligentiescores in groep 4 en de prestaties twee of vier jaar later in groep 6, respectievelijk 8. Voor deze analyses wordt geen gebruik gemaakt van de representatieve, maar van de totale steekproef. De reden daarvoor is dat er bij elk van de vijf metingen sprake is geweest van uitval. Die uitval heeft voornamelijk te maken met het feit dat tussen twee metingen complete scholen niet meer deelnemen (steeds ongeveer een derde) en dat (daarnaast) ook individuele leerlingen niet meer deelnemen (ongeveer een kwart door verandering van school, zittenblijven en verwijzing naar het SBO). De uitval van scholen (die steeds zijn aangevuld met vergelijkbare scholen) is niet selectief, die van leerlingen wel – wat voor de hand ligt vanwege het zittenblijven en verwijzen naar het SBO. Dit laatste houdt in dat de prestaties van de uitvallers lager zijn dan die van de niet-uitvallers (vgl. Roeleveld & Vierke, 2003). De uiteindelijke selectiviteit van de longitudinale groep leerlingen die in groep 8 terechtkomt valt echter mee, namelijk qua toetsprestaties .10 tot .18 standaarddeviatie ten opzichte van de uitvallers (Roeleveld, z.j.). Een consequentie van dit alles is dat de aantallen afhankelijk zijn van het aantal metingen waarover het gaat. Een indicatie is dat het bij

drie metingen (ofwel groep 4, 6 en 8) circa 5000 leerlingen betreft en bij twee metingen (bv. groep 6 en 8) 8000.

## 2.2 Variabelen

Voor de onderhavige analyses is een aantal variabelen geselecteerd uit de PRIMA-bestanden. Hoewel de meeste variabelen bij elk van de vijf metingen op dezelfde wijze zijn geoperationaliseerd, is er bij sommige toch sprake van wijzingen. Hierna volgt een overzicht van de geselecteerde variabelen. Omdat de intelligentietests hier centraal staan, wordt daar uitgebreider op ingegaan dan op de andere variabelen. Wat betreft de afname van de intelligentietests en taal-, reken- en leestoetsen is het van belang er op te wijzen dat deze voor de eerste PRIMA-meting heeft plaatsgevonden in de periode september-december en bij de latere metingen in januari-maart. De consequentie hiervan is dat de scores bij de eerste meting lager liggen dan die bij de latere metingen. Bij de analyse en interpretatie van de ontwikkeling over de vijf metingen heen zal met dit gegeven rekening worden gehouden.

### *Intelligentie*

Bij PRIMA is gekozen voor een intelligentietest die oorspronkelijk door het RION is ontwikkeld ten behoeve van de Landelijke Evaluatie van het Onderwijsvoorrrangsbeleid (LEO) (Dodema-Winsemius & Van der Werf, 1988). Bij LEO is gebruik gemaakt van vijf subtests verdeeld over drie categorieën ‘ruimtelijk’, ‘symbolisch’ en ‘semantisch’; voor groep 4 was de test in zijn geheel non-verbaal, voor de groepen 6 en 8 bestond hij uit een verbaal en non-verbaal deel. Voor PRIMA zijn alleen de twee nonverbale subtests Figuren samenstellen en Exclusie geselecteerd; deze behoren tot de categorie ‘ruimtelijk’.<sup>2</sup> Bij Figuren samenstellen moet een voorgegeven figuur, bijvoorbeeld een vierkant, worden samengesteld uit een reeks onregelmatig gevormde segmenten. Bij Exclusie moet in een reeks figuren er één worden aangewezen dat er niet in past. Voor deze twee – nonverbale – subtests is mede gekozen om eventuele bias tegenover allochtone leerlingen te vermijden (vgl. Extra & Verhoeven, 1985). Overigens was in LEO-onderzoek al geconcludeerd dat er geen aanwijzingen zijn voor testbias. Factoranalyse liet namelijk zien dat de subtests bij allochtone leerlingen hetzelfde meten als bij autochtone leerlingen (Tesser, Mulder & Van der Werf, 1991). Ook is uit ander Nederlands onderzoek naar cultuur-fair toetsen, meer specifiek naar de Eindtoets basisonderwijs van het Cito, niet gebleken dat er sprake is van systematische bias (Uiterwijk, 1994). Ander Nederlands en international onderzoek naar culturele bias bij het testen van intelligentie wijst eveneens op hooguit geringe bias (Te Nijenhuis & Van der Flier, 2003; Te Nijenhuis, De Jong, Evers & Van der Flier, 2004).

Er bestaan drie verschillende versies van de twee subtests, namelijk voor groep 4, 6 en 8. In een vooronderzoek van PRIMA zijn de tests in de verschillende groepen afgenomen en opnieuw op hun psychometrische kwaliteiten onderzocht. Daaruit bleek dat er zich geringe plafondeffecten voordeden en dat de tests minder goed differentieerden tussen kinderen met een verschillende sociaal-etnische achtergrond (m.n. niet zo goed tussen de autochtone en allochtone achterstandsléerlingen, ofwel de 1.25- en 1.90-leerlingen, maar wel tussen deze twee achterstandscategorieën en de niet-achterstandsléerlingen, ofwel de 1.00-leerlingen). Om die redenen zijn enkele

---

2 Deze subtests tonen veel verwantschap met de bijvoorbeeld de RAKIT-onderdelen Closure en Exclusie; vgl. Te Nijenhuis, Tolboom, Resing & Bleichrodt, 2004; Evers, Vliet-Mulder & Groot, 2000a,b).



items uit de oorspronkelijke tests verwijderd, waardoor de testparameters iets gunstiger, maar nog steeds niet optimaal, werden (Driessen, Van Langen & Oudenhoven, 1994). Opgemerkt kan nog worden dat in de (eerdere) LEO-onderzoeken de tests beter differentieerden tussen de (dezelfde) sociaal-etnische leerlingcategorieën.

De tests zijn voor elk van de vijf PRIMA-metingen identiek gebleven. Voor groep 4 telt Figuren samenstellen 17 en Exclusie 20 items; voor groep 6 en 8 gaat het om 19, respectievelijk 15 items. De toets als geheel omvat dus 37, 34 en 34 items. De scores op de (sub)tests zijn uitgedrukt als het aantal goede gemaakte opgaven. Vanwege de verschillen in moeilijkheidsgraad en aantallen items zijn de scores niet zonder meer vergelijkbaar over de groepen heen.

#### *Taal, rekenen en begrijpend lezen*

In de groepen 4, 6 en 8 zijn taal- en rekentoetsen afgenomen en in de groepen 6 en 8 ook toetsen begrijpend lezen. De toetsen verschillen steeds per groep qua moeilijkheidsgraad.

*Taal.* De taaltoetsen zijn in de vijf PRIMA-metingen hetzelfde gebleven. De scores op deze toetsen zijn na een kalibratieprocedure omgezet naar PRIMA-vaardigheidsscores (Vierke, 1995). Dit houdt in dat de scores over de groepen heen op één schaal zijn gebracht en met elkaar vergeleken kunnen worden.

*Rekenen.* Voor de eerste twee PRIMA-metingen zijn de LEO-rekentoetsen gebruikt. Ook deze toetsresultaten zijn omgezet in PRIMA-vaardigheidsscores. Met ingang van de derde PRIMA-meting zijn echter de rekenen/wiskunde-toetsen uit het Cito-Leerlingvolgsysteem afgenomen. De scores op deze toetsen zijn eveneens omgezet in vaardigheidsscores, maar hier betreft het de Cito-vaardigheidsscores. Vergelijking met de eerdere twee metingen is daardoor niet mogelijk.

*Lezen.* Met ingang van de derde PRIMA-meting is in de groepen 6 en 8 ook de begrijpend lezen-toets uit het Cito-Leerlingvolgsysteem afgenomen. De toetsresultaten zijn ook omgezet in Cito-vaardigheidsscores.

#### *Leerlingachtergronden*

Via de scholen zijn een aantal achtergrondkenmerken verzameld; ze hebben betrekking op de leerling zelf en diens ouders en gezinssituatie.

*Sekse.* (1) Jongen, (2) meisje.

*Opleiding.* Hoogst voltooide opleiding van de ouders: (1) maximaal lager onderwijs, (2) maximaal lbo, (3) maximaal mbo, (4) hbo/wo.

*Etniciteit.* Het geboorteland van de ouders: (1) Nederlands, (2) gemengd: Nederland en buitenland, (3) Surinaams/Antilliaans, (4) Turks, (5) Marokkaans, (6) overig buitenland.

### **2.3 Methode**

Het doel van deze rapportage is primair beschrijvend van aard. Om die reden wordt gebruik gemaakt van traditioneel-beschrijvende analysetechnieken zoals variantie- en regressie-analyse. Daarmee wordt vooral inzicht gegeven in de verdelingskenmerken van en samenhangen tussen de relevante variabelen, al dan niet over een aantal jaren heen.

### 3 Resultaten

#### 3.1 Ontwikkelingen in intelligentie?

Eerst is nagegaan of er zich over de vijf PRIMA-metingen heen ontwikkelingen hebben voorgedaan in de scores op de intelligentietests. De centrale vraag luidt hier of de scores op de intelligentietests over de jaren heen stabiel zijn gebleven of dat er sprake is van stijgingen dan wel dalingen. Bijkomende vraag is of er daarbij verschillen zijn die te maken hebben met de achtergronden van de leerlingen, i.c. sekse, milieu, etniciteit. Om deze vragen te beantwoorden is gebruik gemaakt van transversale analyse van de representatieve steekproef, dat wil zeggen dat de gegevens van verschillende, opeenvolgende cohorten over de tijd worden vergeleken. In Tabel 1 staan de resultaten uitgesplitst naar meting en groep. In de cellen staan allereerst het gemiddelde, vervolgens de standaarddeviatie van de totale steekproef, en als laatste de correlatiecoëfficiënt *eta*. Die *eta* is berekend over de laatste vier metingen; dat is gedaan omdat (zoals eerder aangegeven) de eerste meting veel eerder in het schooljaar heeft plaatsgevonden, waardoor de samenhang over alle vijf de metingen een minder adequaat beeld zou geven.

*Tabel 1 – Intelligentiescores naar PRIMA-meting (gemiddelden)*

Groep	Test	PRIMA					SD <sup>a</sup>	<i>eta</i> <sup>b</sup>
		1	2	3	4	5		
4	Figuren	11.8	12.3	12.4	12.2	12.2	3.0	.03
	Exclusie	14.4	14.8	15.7	15.7	15.7	3.6	.11
	Totaal	26.1	27.1	28.2	28.1	27.9	5.3	.08
6	Figuren	13.7	14.1	14.5	14.3	14.3	2.9	.05
	Exclusie	11.1	11.4	11.6	11.6	11.6	2.7	.04
	Totaal	24.7	25.5	26.1	26.0	25.9	4.7	.05
8	Figuren	13.9	14.0	14.3	14.4	14.3	3.1	.03
	Exclusie	11.7	11.8	11.9	12.0	11.9	2.1	.04
	Totaal	25.5	25.8	26.2	26.4	26.2	4.4	.04

<sup>a</sup> SD totaal PRIMA 1 - 5; <sup>b</sup> *eta* over PRIMA 2 - 4

De tabel maakt duidelijk dat er slechts zeer geringe verschillen zijn tussen de jaren. Zoals verwacht liggen de scores van de eerste meting iets lager dan die van volgende metingen. Daarnaast zijn ook de scores van de tweede meting iets lager,<sup>3</sup> maar nogmaals: de verschillen zijn over het geheel genomen minimaal (zowel in relatie tot de standaarddeviatie als op basis van de *eta*).

<sup>3</sup> Eerder werd er eveneens een lichte stijging tussen de tweede en derde PRIMA-meting geconstateerd met betrekking tot de toetsprestaties voorbereidend lezen en rekenen in groep 2. Deze stijging was echter ook zeer gering en zette zich daarna ook niet meer door. Een verklaring voor deze ontwikkeling werd overigens niet gevonden (vgl. Driessen, 2005; Roeleveld, 2003).

Dat er over de metingen heen geen verschillen zijn qua intelligentiescores, wil niet zeggen dat er tussen subgroepen geen verschillen zouden kunnen bestaan. Daarom zijn bovenstaande analyses herhaald, maar nu met een verdere uitsplitsing naar sekse, milieu (i.c. opleiding ouders) en etniciteit. Er is hier dus nagegaan of bijvoorbeeld de scores van de jongens zich over de metingen heen anders hebben ontwikkeld dan die van de meisjes (hier is nog niet nagegaan of er verschillen zijn tussen de scores van jongens en meisjes; dat volgt later).

Wat betreft de sekse van de leerlingen is het patroon van de scores van jongens, respectievelijk meisjes identiek aan het algemene patroon zoals dat in Tabel 1 is weergegeven. De verschillen tussen de metingen bedragen doorgaans hooguit enkele tienden punten. Om die reden is hier afgezien van opname van uitgebreide reeksen tabellen. Dezelfde analyses maar nu met milieu en etniciteit leveren een vergelijkbaar resultaat op. Met andere woorden: er zijn binnen elk van de categorieën van de achtergrondkenmerken sekse, milieu en etniciteit geen verschillen tussen de PRIMA-metingen wat betreft de scores op de intelligentietests anders dan die ook voor de hele steekproef gelden.

In een volgende serie analyses is nagegaan of er qua intelligentie binnen elk van de metingen verschillen zijn tussen de categorieën van de achtergrondkenmerken, bijvoorbeeld of er verschillen zijn tussen jongens en meisjes. Omdat de resultaten voor elk van de metingen tot vergelijkbare resultaten leidden, worden hier alleen de bevindingen met betrekking tot de vijfde meting gepresenteerd; ze staan in Tabel 2.

Tabel 2 – Intelligentiescores naar achtergrondkenmerken; PRIMA-5 (gemiddelden)

Seks								part. <i>eta</i>	
Groep	Test	Jongen	Meisje	<i>eta</i>					
4	Figuren	12.4	12.0	.06				.06	
	Exclusie	15.3	16.2	.13				.13	
	Totaal	27.7	28.2	.05				.05	
6	Figuren	14.4	14.3	.01				.01	
	Exclusie	11.4	11.8	.08				.08	
	Totaal	25.7	26.1	.05				.05	
8	Figuren	14.2	14.3	.02				.02	
	Exclusie	11.8	12.0	.03				.03	
	Totaal	26.1	26.3	.03				.06	
Opleiding								part. <i>eta</i>	
	Test	LO	LBO	MBO	HBO/WO	<i>eta</i>	<i>r</i>		
4	Figuren	10.8	11.6	12.2	12.9	.20	.20	.16	
	Exclusie	14.0	14.8	15.9	16.5	.21	.21	.17	
	Totaal	24.8	26.4	28.1	29.4	.25	.25	.20	
6	Figuren	13.1	13.7	14.3	14.9	.19	.19	.15	
	Exclusie	10.5	11.0	11.7	12.3	.22	.22	.19	
	Totaal	23.6	24.7	26.1	27.2	.24	.24	.20	
8	Figuren	13.3	13.6	14.3	15.1	.19	.19	.18	
	Exclusie	11.5	11.5	12.0	12.3	.16	.15	.16	
	Totaal	24.8	25.2	26.3	27.4	.21	.21	.20	
Etniciteit								part. <i>eta</i>	
	Test	Ned	Gemengd	Sur/Ant	Tur	Mar	Overig	<i>eta</i>	
4	Figuren	12.4	12.4	10.9	11.1	10.7	11.2	.17	.11
	Exclusie	16.0	15.7	14.3	14.3	13.7	15.2	.17	.10
	Totaal	28.4	28.1	25.2	25.4	24.4	26.4	.20	.12
6	Figuren	14.4	14.6	12.8	13.4	12.8	13.8	.16	.11
	Exclusie	11.8	11.7	10.3	10.7	10.5	11.3	.15	.10
	Totaal	26.2	26.3	23.1	24.1	23.3	25.1	.18	.12
8	Figuren	14.4	14.6	13.1	13.8	13.3	13.6	.11	.06
	Exclusie	12.0	11.9	11.4	11.7	11.5	11.9	.07	.03
	Totaal	26.4	26.5	24.5	25.5	24.9	25.5	.11	.06

Uit het bovenste deel van de tabel volgt dat er geen verschillen zijn tussen jongens en meisjes. Duidelijke verschillen zijn er wel met betrekking tot de opleiding van de ouders: er is een monotoon verband tussen de ouderlijke opleiding en de intelligentiescores van hun kinderen. Vergelijking van de *eta* en Pearson *r* in de betreffende kolommen van de tabel laat bovendien zien dat deze samenhang perfect lineair is. Er is eveneens een duidelijke relatie met etniciteit, al is deze samenhang wat zwakker dan voor opleiding ouders. De scores van de kinderen uit autochtone en gemengd autochtone-allochtone gezinnen zijn het hoogste en nagenoeg gelijk. In groep 4 zijn de scores van de Marokkaanse kinderen het laagst, in groep 6 en 8 zijn dat die van de

Surinaamse en Antilliaanse kinderen. Zoals al opgemerkt geldt het bovenstaande voor elk van de metingen in ongeveer gelijke mate.

### **3.2 Intelligentie en achtergrondkenmerken**

Een belangrijke vraag is in hoeverre intelligentieverschillen verklaard kunnen worden door (het samenspel van) sekse, milieu en etniciteit. In de laatste kolom van Tabel 2 staan de via variantie-analyse verkregen partiële *eta*'s (vgl. *beta*'s). Die coëfficiënten geven een indruk van het relatieve gewicht van de drie achtergrondkenmerken in de verklaring van verschillen in intelligentiescores. Het blijkt dat het effect van opleiding na correctie voor sekse en etniciteit doorgaans wat kleiner wordt, maar in alle gevallen toch het sterkste blijft. Ook is nagegaan of er zich tweede en derde orde interactie-effecten voordoen, bijvoorbeeld sekse x etniciteit. Daar blijkt echter geen sprake van te zijn. Dit betekent dus bijvoorbeeld dat er ook binnen elk van de etnische groepen geen verschillen zijn in intelligentie tussen jongens en meisjes.

### **3.3 Intelligentie, prestaties en achtergrondkenmerken**

In Tabel 3 staan de correlaties tussen de intelligentietestcores, toetsprestaties en achtergrondkenmerken (waarvoor etniciteit is uiteengelegd in dummy-variabelen). De samenhangen met de achtergrondkenmerken staan overigens deels ook al in Tabel 2. Het betreft ook hier gegevens van de vijfde PRIMA-meting. De correlaties voor de eerdere metingen zijn – ondanks verschillen in rekentoetsen – nagenoeg gelijk; ze verschillen hooguit enkele honderdste punten.<sup>4</sup>

---

<sup>4</sup> Gememoreerd moet worden dat zowel de intelligentietests als prestatietoetsen in de drie jaargroepen weliswaar dezelfde concepten meten, maar dat ze verschillen qua moeilijkheidsgraad.

*Tabel 3 – Samenhangen intelligentie, prestaties en achtergrondkenmerken, naar groep; PRIMA-5 (correlaties)*

Groep	Test	Intelligentie		Prestaties					Achtergrondkenmerken					
		Exc	Tot	Tl	Rek	Lez	Sek	Opl	Ned	Gem	S/A	Tur	Mar	Ov
4	Fig	.35	.78	.28	.41	-	-.06	.20	.13	.02	-.06	-.07	-.10	-.08
	Exc		.86	.30	.35	-	.13	.21	.14	-.00	-.06	-.08	-.11	-.04
	Tot			.35	.46	-	.05	.25	.17	.01	-.08	-.10	-.13	-.07
6	Fig	.44	.86	.29	.40	.31	.01	.19	.12	.03	-.08	-.06	-.10	-.04
	Exc		.83	.34	.42	.41	.08	.22	.12	.01	-.08	-.08	-.09	-.03
	Tot			.37	.48	.42	.05	.24	.14	.02	-.09	-.08	-.11	-.04
8	Fig	.43	.91	.32	.46	.37	.02	.19	.08	.02	-.06	-.04	-.06	-.05
	Exc		.77	.27	.33	.31	.03	.15	.05	-.00	-.04	-.03	-.04	-.01
	Tot			.35	.48	.41	.03	.21	.08	.01	-.06	-.04	-.06	-.04

De samenhangen tussen de intelligentiescores en de rekenvaardigheid zijn steeds het hoogste, i.c. matig sterk tot sterk. Dit is niet zo verwonderlijk, omdat de intelligentietests non-verbaal van aard zijn. De samenhangen met begrijpend lezen zijn wat zwakker en die met de taalvaardigheid nog weer zwakker, i.c. doorgaans matig sterk.

De vraag is vervolgens of die samenhangen voor elk van de onderscheiden etnische groepen even sterk zijn. Om daar zicht op te krijgen zijn de correlationele analyses waarvan de resultaten in Tabel 3 zijn gepresenteerd herhaald voor elk van de zes etnische groepen apart. Daaruit bleek een zeer grote mate van overeenstemming. Voorzover er zich afwijkingen voordoen van het algemene beeld, dienen die waarschijnlijk te worden toegeschreven aan de verlaging van de variantie doordat de afzonderlijke etnische groepen homogener zijn samengesteld dan de totale steekproef (vgl. Sirin, 2005).

### 3.4 De longitudinale ontwikkeling van intelligentie

Een vraag luidt hoe de intelligentiescores zich in de opeenvolgende jaren tot elkaar verhouden. De vraag is dan of leerlingen die in het eerste jaar (c.q. groep) hoog scoren dat ook in latere jaren (c.q. groepen) nog blijven doen. Anders gezegd: hoe stabiel scoren de leerlingen op de dimensies die worden gemeten met de intelligentietests? Voor deze – en volgende – analyses is de totale PRIMA-steekproef genomen. In Tabel 4 staan de correlaties tussen de scores in groep 4, 6 en 8, waarbij groep 8 betrekking heeft op respectievelijk de derde, vierde en vijfde PRIMA-meting. De labels geven de combinatie van (sub)test, PRIMA-meting en groep weer; Fig 14 wil bijvoorbeeld zeggen: subtest Figuren, PRIMA 1, groep 4.

*Tabel 4 – Samenhangen intelligentie groep 4, 6 en 8, longitudinaal (correlaties)*

	Fig 26	Fig 38	Exc 26	Exc 38	Tot 26	Tot 38
Fig 14	.47	.39				
Exc 14			.35	.19		
Tot 14					.52	.39
	Fig 36	Fig 48	Exc 36	Exc 48	Tot 36	Tot 48
Fig 24	.49	.44				
Exc 24			.31	.18		
Tot 24					.51	.41
	Fig 46	Fig 58	Exc 46	Exc 58	Tot 46	Tot 58
Fig 34	.45	.47				
Exc 34			.29	.22		
Tot 34					.48	.45

De tabel laat zien dat de samenhangen tussen de scores in groep 4 en groep 6 en 8 voor de subtest Figuren en de totale test matig sterk tot sterk zijn, waarbij de samenhangen tussen groep 4 en 6 in het algemeen wat sterker zijn dan die met groep 8. De samenhangen voor de subtest Exclusie daarentegen zijn opvallend lager, voor groep 4 met groep 6 matig sterk en met groep 8 tussen zwak en matig sterk in. Dit betekent dus dat leerlingen die in groep 4 hoog scoren op Exclusie in latere groepen niet noodzakelijkerwijs ook hoog scoren op die dimensie. De stabiliteit van de tests is daarmee dus vrij laag en de scores geven voornamelijk een momentopname weer.

Deze analyses zijn vervolgens herhaald voor elk van de categorieën van sekse, milieu en etniciteit. Vanwege de eenvormigheid in resultaten wordt afgezien van presentatie van reeksen van tabellen. Het algemene beeld is dat de sterkte van de correlaties en trends daarin doorgaans ook voor elk van de onderscheiden categorieën van de drie achtergrondkenmerken geldt.

### 3.5 De voorspelling van de prestaties

In een volgende reeks van analyses is nagegaan in welke mate de prestaties van de leerlingen in groep 8 voorspeld kunnen worden. In het bijzonder gaat het daarbij om de vraag wat de bijdrage is van intelligentie in die voorspelling, ofwel: voegen de intelligentiescores nog iets toe bovenop de achtergrondkenmerken en prestaties van de leerlingen?<sup>5</sup> Voor deze regressie-analyses zijn eerst een aantal uiteenlopende modellen beproefd; uiteindelijk is geopteerd voor de volgende opzet.

De te verklaren variabelen zijn de prestaties op respectievelijk taal, rekenen en begrijpend lezen in groep 8. In model A worden alleen de achtergrondkenmerken in de analyse ingevoerd: sekse (1=jongen; 0=meisje), opleiding (1=lo, 2=lbo, 3=mbo, 4=hbo/wo), en etniciteit (1=Nederlands plus gemengd; 0=overig). In model P worden de prestaties in groep 4 en 6 opgenomen. In Mo-

<sup>5</sup> Deels gaat het hier ook om de vraag naar de predictieve validiteit van de intelligentietests.



del I gaat het om de scores op de intelligentietests (Figuren en Exclusie apart) in groep 4, 6 en 8. In model A+P worden de achtergrondkenmerken plus de prestaties ingevoerd en in model I+P de prestaties plus intelligentie. In model A+P+I worden de drie modellen gecombineerd. In de tabel staan de gestandaardiseerde coëfficiënten (*beta*'s); deze geven een indicatie van het relatieve belang van de voorspellers in het betreffende model. Onderin de tabel staan de percentages verklaarde variantie ( $%R^2$ ). Tabel 5 geeft de gegevens met betrekking tot de taalprestaties.

*Tabel 5 – Regressie-analyse taalprestaties groep 8 PRIMA-5 (beta's)*

	Model					
	A	P	I	A+P	I+P	A+P+I
Sekse	-.02			.02		.03
Opleiding	.28			.11		.09
Etniciteit	.17			.01		.02
Taal 4		.20		.17	.17	.14
Taal 6		.54		.51	.49	.46
Fig 4			.06		.01	.00
Exc 4			.17		.06	.06
Fig 6			.04		.00	.00
Exc 6			.15		.06	.06
Fig 8			.15		.11	.11
Exc 8			.07		.04	.03
% $R^2$	15.6	43.3	18.0	44.6	46.3	47.3

Uit de tabel volgt onder model A dat van de drie achtergrondkenmerken de opleiding van de ouders het sterkste effect heeft, gevolgd door hun etniciteit. Sekse hangt niet samen de taalprestaties.<sup>6</sup> Samen verklaren de achtergrondkenmerken 15.6% van de verschillen in taalprestaties. Bij de verklaring van de taalprestaties door de toetsscores in groep 4 en 6 (Model P) domineren die van groep 6 sterk. De beide eerdere prestaties verklaren aanzienlijk meer dan de achtergrondkenmerken, bijna het drievoudige. Het model met alleen de intelligentiescores (Model I) laat hooguit zwakke effecten zien; samen binden de zes tests 18% van de variantie. Als de achtergrondkenmerken en eerdere prestaties tegelijk worden ingevoerd (Model A+P), blijft van de effecten van de achtergrondkenmerken weinig meer over; er resteert alleen nog een zwak effect van opleiding. Uit vergelijking van de percentages verklaarde variantie kan ook worden afgeleid dat de achtergrondkenmerken slechts 1% verklaarde variantie toevoegen aan de prestaties. Het model met prestaties en intelligentie (Model I+P) laat zien dat intelligentie 3% toevoegt aan de verklaring door eerdere prestaties. Er is daarbij alleen van Figuren groep 8 een zeer klein effect. Uit het model waarin ook de intelligentiescores zijn ingevoerd (Model A+P+I) blijkt dat er alleen van Figuren groep 8 een zwak effect uitgaat op de taalprestaties. Samen verklaren de zes intelligentietests minder dan 3% extra variantie.

<sup>6</sup> De hoogte van coëfficiënten wijkt af van die uit Tabel 2. Dat heeft te maken met het feit dat hier een andere (longitudinale en met meer variabelen uitgebreide) steekproef is geanalyseerd en etniciteit als een dichotome variabele is geoperationaliseerd.

In Tabel 6 staan de resultaten met betrekking tot de rekenprestaties en in Tabel 7 met betrekking tot de leesprestaties.

*Tabel 6 – Regressie-analyse rekenprestaties groep 8 PRIMA-5 (beta's)*

	Model					
	A	P	I	A+P	I+P	A+P+I
Sekse	.16			.04		.06
Opleiding	.30			.10		.09
Etniciteit	.00			-.06		-.06
Rekenen 4		.16		.16	.16	.15
Rekenen 6		.69		.67	.63	.61
Fig 4			.11		.01	-.03
Exc 4			.11		-.03	.02
Fig 6			.10		.00	.01
Exc 6			.15		.02	.02
Fig 8			.23		.13	.13
Exc 8			.10		.05	.04
% $R^2$	11.8	66.0	30.3	66.9	67.9	68.9

Ten aanzien van de rekenprestaties laat Tabel 6 zien dat er geen effect is van etniciteit, een matig sterk effect is van opleiding en een zwak effect van sekse. Dit laatste is niet onverwacht: jongens zijn beter qua rekenen dan meisjes. De eerdere rekenprestaties verklaren een zeer groot deel (tweederde) van de latere prestaties. De intelligentietests verklaren bijna een derde van de variantie. Dit is vrij veel, maar kan worden toegeschreven aan het feit dat non-verbale intelligentie sterker zal samenhangen met rekenen dan taal. De combinatie van intelligentie en eerdere prestaties verklaart minder dan 2% bovenop de prestaties; alleen van Figuren groep 8 gaat nog een zwak effect uit. Het laatste model maakt duidelijk dat zowel de achtergrondkenmerken als ook de intelligentiescores praktisch niets toevoegen aan de eerdere prestaties, ten opzichte van het model met alleen de prestaties minder dan 3%. Daarbij dient overigens bedacht te worden dat het om zes subtests gaat.

Tabel 7 – Regressie-analyse leesprestaties groep 8 PRIMA-5 (beta's)<sup>7</sup>

	Model					
	A	P	I	A+P	I+P	A+P+I
Sekse	-.06			-.01		-.01
Opleiding	.34			.13		.11
Etniciteit	.05			-.04		-.04
Lezen 6		.74		.70	.66	.63
Fig 4			.02		-.02	-.02
Exc 4			.18		.07	.07
Fig 6			.06		-.00	.00
Exc 6			.15		.02	.02
Fig 8			.18		.13	.12
Exc 8			.10		.06	.06
% $R^2$	13.5	54.3	22.6	55.5	57.4	58.3

De resultaten ten aanzien van de leesprestaties vertonen veel overeenkomsten met die van de taalprestaties; vandaar dat wordt afgezien van een uitvoerige bespreking.

Als we nu de uitkomsten van de drie analyses met elkaar vergelijken kan het volgende worden geconcludeerd. De eerdere prestaties verklaren in alle modellen verreweg het grootste deel van de verschillen in de prestaties van groep 8. Van de achtergrondkenmerken is de opleiding van de ouders de belangrijkste predictor. Bij de taalprestaties is er nog een zwak effect van etniciteit, en bij de rekenprestaties van sekse. De intelligentietests verklaren binnen de drie prestatiedomeinen het meeste van de rekenprestaties. In de modellen met zowel achtergrondkenmerken, prestaties als intelligentie, is de extra verklaaringskracht van intelligentie zeer gering, bij de taal- en leesprestaties 4% en bij de rekenprestaties 3% bovenop de verklaring van de eerdere prestaties. In deze modellen is er alleen van de test Figuren groep 8 een consistent (zwak) effect.

Uit het bovenstaande kan worden afgeleid dat de extra verklaaringskracht van intelligentie maar zeer gering is. In de betreffende modellen zijn de effecten van de zes in verschillende groepen afgenomen subtests voor elkaar gecorrigeerd. Een vervolgvraag luidt of intelligentietests die de groepen 6 en 8 zijn afgenomen nog iets toevoegen aan de tests die in groep zijn afgenomen. Om daar zicht op te krijgen zijn de analyses voor Model A+P+I herhaald, maar met als verschil dat nu nadat eerst de achtergrondkenmerken en prestaties zijn ingevoerd de intelligentietests per groep stapsgewijs zijn ingevoerd. In Tabel 8 staan de percentages verklaarde variantie die per stap worden toegevoegd weergegeven.

<sup>7</sup> In groep 4 is geen leestoets afgenomen.

*Tabel 8 – Toegevoegde percentages verklaarde variantie; taal- reken- en leesprestaties groep 8 (PRIMA-5)*

	Taal	Prestaties	
		Rekenen	Lezen
Achtergronden	15.6	11.8	13.5
Eerdere prestaties	29.0	55.2	42.0
Intelligentie 4	1.2	.2	1.0
Intelligentie 6	.7	.4	.4
Intelligentie 8	.1	1.4	1.4
Totaal	47.3	68.9	58.3

Uit de tabel volgt dat intelligentie voor taal en rekenen 2% extra variantie verklaart bovenop achtergronden en prestaties, en voor lezen 3%. Voor taal verklaren de intelligentietests in groep 4 het meest, en voegen die van groep 6 en 8 nauwelijks nog iets toe. Voor rekenen is het omgekeerde het geval; daar voegen de tests van groep 8 nog wat toe. En voor lezen geldt weer een andere verdeling: de tests uit groep 8 en 4 voegen het meeste toe.

In Tabel 9 wordt nog op een andere manier duidelijk gemaakt wat de betekenis is van intelligentie als controlevariabele. In die tabel staan de afwijkingsscores ten opzichte van het gemiddelde voor achtereenvolgens taal, rekenen en lezen. Eerst de ongecorrigeerde scores per categorie van sekse, opleiding en etniciteit, dan de scores waarbij is gecorrigeerd voor de invloed van de andere twee achtergrondkenmerken, en als laatste de scores waarbij eerst is gecorrigeerd voor de intelligentiescores. Om de afwijkingsscores adequaat te kunnen plaatsen zijn onder in de tabel per prestatietoets ook het algemeen gemiddelde en de standaarddeviatie opgenomen.

*Tabel 9 – Afwijkingsscores taal- reken- en leesprestaties groep 8; ongecorrigeerd, gecorrigeerd voor achtergrondkenmerken, gecorrigeerd voor achtergrondkenmerken en intelligentie (PRIMA-5)*

		Prestaties								
		Taal			Rekenen			Lezen		
		Ong	A	A+I	Ong	A	A+I	Ong	A	A+I
Sekse	jongen	-1	-1	-0	1	1	2	-1	-1	-1
	meisje	1	1	0	-2	-2	-2	1	1	1
Opleiding	lo	-21	-12	-9	-4	-3	-3	-8	-6	-5
	lbo	-9	-9	-6	-3	-3	-2	-5	-5	-4
	mbo	5	3	2	1	1	1	2	1	1
	hbo/wo	18	15	11	4	3	2	9	8	6
Etniciteit	Ned	6	4	3	1	0	-0	2	1	0
	gemengd	7	4	4	0	-0	-1	3	1	1
	Sur/Ant	-10	-6	-4	-2	-1	-1	-4	-2	-2
	Tur	-27	-20	-18	-2	-1	0	-9	-5	-4
	Mar	-17	-8	-5	-3	-1	0	-6	-1	-0
	overig	-10	-7	-7	-0	1	1	-1	1	1
Gemiddelde			1116			117			54	
SD			35			9			16	

Met betrekking tot de taalprestaties zien we onder de ongecorrigeerde conditie (kolom 'Ong') dat jongens 1 punt lager scoren dan meisjes, en kinderen van ouders met hooguit lager onderwijs 21 punten onder het algemeen gemiddelde (van 1116) scoren. Kinderen van Turkse ouders scoren 27 punten onder het algemeen gemiddelde. Als gecorrigeerd wordt voor de overige achtergrondkenmerken (kolom 'A'), verbeteren de scores van kinderen van laagopgeleide en allochtone ouders. Als vervolgens ook nog wordt gecorrigeerd voor intelligentie (kolom 'A+I'), dan stijgen de scores van de laagopgeleide en allochtone ouders opnieuw, maar het verschil met de niet voor intelligentie gecorrigeerde conditie is slechts zeer gering (maximaal een tiende standaarddeviatie). Met andere woorden: de functie van intelligentie als controlevariabele is verwaarloosbaar.



## 4 Conclusies

Over de vijf PRIMA-metingen heen zijn er geen verschillen in de hoogte van de scores op de intelligentietests, en dat geldt tevens voor de subtests Figuren en Exclusie, voor de jaargroepen 4, 6 en 8, en voor jongens en meisjes, de onderscheiden ouderlijke opleidingsniveaus en etnische groepen.

Bij de verklaring van verschillen in intelligentiescores – waarbij intelligentie dus als een effectvariabele wordt opgevat – weegt het ouderlijke opleidingsniveau het zwaarste mee, de etnische herkomst is minder relevant, en sekse is helemaal niet van belang.

De intelligentiescores hangen het sterkste samen met de rekenvaardigheid, wat gezien het non-verbale karakter van de tests niet verwondert. De samenhang met begrijpend lezen is minder sterk en met de taalvaardigheid nog weer zwakker. Deze trend doet zich voor binnen elk van de onderscheiden etnische groepen.

De samenhangen tussen de intelligentiescores in lagere en hogere jaren zijn voor de subtest Figuren matig sterk tot sterk, voor Exclusie zijn ze echter beduidend lager. Dit geldt voor elk van de categorieën van de achtergrondkenmerken. De stabiliteit van de scores over de jaargroepen heen is daarmee vrij laag; ze dienen waarschijnlijk vooral te worden opgevat als een momentopname.

In de voorspelling van de taal-, reken- en leesprestaties in groep 8 heeft intelligentie een substantiële, maar per prestatiedomein verschillende bijdrage; voor rekenen is die het grootst. Echter, als controlevariabele is de betekenis van intelligentie minimaal. Extra controle voor intelligentie, bovenop die van achtergrondkenmerken en eerdere prestaties, leidt slechts tot een zeer geringe verbetering van de verklaring van prestatieverschillen.





## Referenties

- Batenburg, T. van, & Werf, M. van der (2004). *NSCCT. Niet Schoolse Cognitieve Capaciteiten Test. Voor groep 4, 6 en 8 van het basisonderwijs. Verantwoording, normering en handleiding*. Groningen: GION.
- Batenburg, T. van, & Werf, G. van der (2005). Second opinion voor de leerkracht. *Didaktief*, 35, (3), 18-19.
- Bosker, R., & Meijnen, G. (2004). *Naar een geïntegreerde opzet van de onderwijscohorten*. Groningen/Amsterdam: GION.
- Ceci, S. (1991). How much does schooling influence general intelligence and its cognitive components. A reassessment of the evidence. *Developmental Psychology*, 27, (5), 703-722.
- Doddema-Winsemius, H., & Werf, M. van der (1988). *Selectie/constructie van toetsen voor sociale redzaamheid en intelligentie ten behoeve van de evaluatie van het OVB*. Groningen: RION.
- Dousma, T., & Horsten, A. (1980). *Tentamineren*. Utrecht/Antwerpen: Het Spectrum.
- Driessen, G. (2002). School composition and achievement in primary education: A large-scale multilevel approach. *Studies in Educational Evaluation*, 28, (4), 347-368.
- Driessen, G. (2005). Cognitieve competenties van jonge kinderen en structurele en culturele gezinskenmerken: Ontwikkelingen en samenhangen. *Pedagogiek*, 25, (2), 103-119.
- Driessen, G., Langen, A. van, & Oudenhoven, D. (1994). *De toetsen voor de cohort Primair Onderwijs. Verantwoording*. Nijmegen: ITS.
- Driessen, G., Langen, A. van, Portengen, R., & Vierke, H. (1998). *Basisonderwijs: Veldwerkverslag, leerlinggegevens en oudervragenlijsten. Basisrapportage PRIMA-cohortonderzoek. Tweede meting 1996-1997*. Nijmegen: ITS.
- Driessen, G., Langen, A. van, & Vierke, H. (2000). *Basisonderwijs: Veldwerkverslag, leerlinggegevens en oudervragenlijsten. Basisrapportage PRIMA-cohortonderzoek. Derde meting 1998/99*. Nijmegen: ITS.
- Driessen, G., Langen, A. van, & Vierke, H. (2002). *Basisonderwijs: Veldwerkverslag, leerlinggegevens en oudervragenlijsten. Basisrapportage PRIMA-cohortonderzoek. Vierde meting 2000/01*. Nijmegen: ITS.
- Driessen, G., Langen, A. van, & Vierke, H. (2004). *Basisonderwijs: Veldwerkverslag, leerlinggegevens en oudervragenlijsten. Basisrapportage PRIMA-cohortonderzoek. Vijfde meting 2002/03*. Nijmegen: ITS.
- Driessen, G., & Werf, G. van der (1992). *Het functioneren van het voorgezet onderwijs. De positie van het leerlingen in het eerste leerjaar*. Nijmegen/Groningen: ITS/GION.
- Extra, G., & Verhoeven, L. (1985). Bias in intelligentie-onderzoek bij allochtone kinderen. *Pedagogische Studiën*, 62, 391-395.
- Evers, A., Vliet-Mulder, J., & Laak, J. ter C. (1992). *Documentatie van tests en testresearch in Nederland*. Assen/Maastricht: Van Gorcum.

- Evers, A., Vliet-Mulder, J., & Groot, C. (2000a). *Documentatie van tests en testresearch in Nederland. Deel I. Testbeschrijvingen*. Assen: Van Gorcum.
- Evers, A., Vliet-Mulder, J., & Groot, C. (2000b). *Documentatie van tests en testresearch in Nederland. Deel II. Testresearch*. Assen: Van Gorcum.
- Friesen, G., & Voestermans, P. (1995). De Bell Curve hype. *De Psycholoog*, 30, (11), 467-470.
- Gennip, H. van (1993). *Instrumenten voor leerlingmetingen in leerjaar 2. Voorstudie cohort-onderzoek primair onderwijs*. Nijmegen: ITS.
- Helms-Lorenz, M., Vijver, F. van de, & Poortinga, Y. (2003). Cross-cultural differences in cognitive performance and Spearman's hypothesis: *g* or *c*? *Intelligence*, 31, 9-29.
- Herrnstein, R., & Murray, C. (1994). *The bell curve. Intelligence and class structure in American life*. New York: The Free Press.
- Hustinx, P., Kuyper, H., & Werf, M. van der (2005). *De onderwijsresultaten van VOCL'89 en VOCL'93 leerlingen verklaard..* Groningen: GION.
- Jensen, A. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 39, (9), 1-123.
- Jungbluth, P., Langen, A. van, Peetsma, T., & Vierke, H. (1996). *Leerlinggegevens basisonderwijs en speciaal onderwijs. Technische rapportage PRIMA-cohortonderzoek 1994/95*. Amsterdam/Nijmegen: SCO/ITS.
- Langen, A. van, Vierke, H., & Robijns, M. (1996). *Veldwerkverslag basisonderwijs en speciaal onderwijs. Technische rapportage PRIMA-cohortonderzoek 1994/95*. Amsterdam/Nijmegen: SCO/ITS.
- Luyten, H., & Bosker, R. (2004). Hoe meritocratisch zijn schooladviezen? *Pedagogische Studiën*, 81, (2), 89-103.
- Mulder, L. (1994). *De derde fase van de OVB-cohortonderzoeken in het basisonderwijs. Het leerlingenonderzoek in schooljaar 1992/1993*. Nijmegen: ITS.
- Mulder, L., m.m.v. P. Tesser & H. Vierke (1993). *De tweede fase van de OVB-cohortonderzoeken in het basisonderwijs. Het leerlingenonderzoek in schooljaar 1990/1991*. Nijmegen: ITS.
- Mulder, L., Vierke, H., & Petersen, B. (1997). *De vierde fase van de OVB-cohortonderzoeken in het basisonderwijs. Het leerlingen, school- en klasonderzoek in schooljaar 1994/1994*. Nijmegen: ITS.
- Neisser, U. e.a. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, (2), 77-101.
- Nijenhuis, J. te, & Flier, H. van der (2003). Immigrant-majority group differences in cognitive performance: Jensen effects, cultural effects, or both? *Intelligence*, 31, 443-459.
- Nijenhuis, J. te, Jong, M.-J. de, Evers, A., & Flier, H. van der (2004). Are cognitive differences between immigrant and majority groups diminishing? *European Journal of Personality*, 18, 405-434.

- Nijenhuis, J. te, Tolboom, E., Resing, W., & Bleichrodt, N. (2004). Does cultural background influence the intellectual performance of children from immigrant groups? *European Journal of Psychological Assessment*, 20, (1), 10-26.
- Roeleveld, J. (2001). *De Cito eindtoets basisonderwijs binnen Prima van 1995 tot en met 2001*. Amsterdam: SCO-Kohnstamm Instituut.
- Roeleveld, J. (2003). *Veranderingen in het kleuteronderwijs*. Amsterdam: SCO-Kohnstamm Instituut.
- Roeleveld, J. (z.j.). *Herkomstenmerken en begintoeets. Secundaire analyses op het PRIMA-cohortonderzoek*. Amsterdam: SCO-Kohnstamm Instituut.
- Roeleveld, J., & Vierke, H. (2003). *Uitval en instroom bij de derde meting van het Prima-onderzoek*. Amsterdam/Nijmegen: SCO-Kohnstamm Instituut/ITS.
- Sirin, S. (2005). Socioeconomic status and academic achievement: A meta-analytical review of research, *Review of Educational Research*, 75, (3), 417-453.
- Tesser, P., Mulder, L., & Werf, G. van der (1991). *De eerste fase van de longitudinale OVB-onderzoeken. Het leerlingonderzoek*. Nijmegen/Groningen: ITS/RION.
- Uiterwijk, J. (1994). *De bruikbaarheid van de Eindtoets Basisonderwijs voor allochtone leerlingen*. Arnhem: CITO.
- Veldhuijzen, N., Goldebeld, P., Sanders, P. (1993). Klassieke testtheorie en generaliseerbaarheidstheorie. In T. Eggen & P. Sanders (Eds.), *Psychometrie in de praktijk* (pp. 33-82). Arnhem: Cito.
- Vierke, H. (1995). *De PRIMA-toetsen gekalibreerd. De ontwikkeling van vaardigheidsscores over de leerjaren heen op basis van de jaargroepoetsen in het cohort primair onderwijs (PRIMA)*. Nijmegen: ITS.



## **Bijlage**

### **Overwegingen met betrekking tot de opname van een intelligentietest in het vervolg op PRIMA**

Met de zesde meting wordt het PRIMA-tijdperk afgesloten. Als vervolg daarop staat een omvattendere cohortstudie gepland, waarbij ook het voortgezet, middelbaar beroeps- en hoger onderwijs wordt betrokken. Deze notitie gaat in op de vraag naar de positie van een intelligentietest in dit nieuwe cohort.

#### **1. De intelligentietest in PRIMA**

Op basis van het hetgeen in het begin van deze notitie is onderzocht, kan worden gesteld dat de PRIMA-intelligentietests bij gebruik als effectvariabele zich gedragen als de prestatie-maten. Opgemerkt dient ook te worden dat de scores op deze tests niet erg stabiel zijn over de jaargroepen heen. Bij gebruik als controlevariabele voegen ze nauwelijks wat toe aan de voorspelling van de prestaties, nadat eerst rekening is gehouden met verschillen in achtergrondkenmerken en eerdere prestaties. Al met al geeft dit dus enige twijfels over de opname van (deze) intelligentietests in het nieuwe cohortonderzoek.

#### **2. De intelligentietest in VOCL**

In de in het voortgezet onderwijs verrichte cohortonderzoeken VOCL'89 en VOCL'93 zijn (eveneens) twee non-verbale tests afgenomen, de PSB-3 en de PSB-8 (Horn, 1969). De eerste test meet de denkvaardigheid of het redeneervermogen ('reasoning'), de tweede het denkbeeldig abstraheren ('abstraction') (Driessen & Van der Werf, 1992). Deze tests zijn vooral gericht op de 'fluid intelligence', dat wil zeggen op de in aanleg aanwezige capaciteiten, los van de door leerervaringen opgedane ervaringen. Beide tests tellen 40 items; de afnametijd bedraagt 8 minuten. In VOCL'89 zijn de tests in leerjaar 1 en 3 afgenomen, in VOCL'93 alleen nog in leerjaar 3. In analyses op de VOCL-data worden de variabelen chronologisch geordend (d.w.z. het moment van afname bepaalt de positie in het verklaringsmodel) en niet theoretisch-causaal (Hustinx, Kuiper & Van der Werf, 2005). Er wordt namelijk vanuit gegaan dat de intelligentiemeting een momentopname is en niet met zekerheid kan worden gezegd dat de uitslag van de test een valide indicator is voor een theoretisch te veronderstellen algemene intellectuele aanleg die ook al eerder in gelijke mate aanwezig was. De correlaties met de loopbaanvariabelen advies, prestatietoetsen en onderwijspositie liggen in leerjaar 1 rond de .27 en in leerjaar 3 rond de .38, en zijn daarmee erg laag. De betreffende onderzoekers hebben twijfels over de validiteit van deze tests, die ze mede toeschrijven aan de beperkte afnameduur en het geringe aantal domeinen dat wordt gedekt. Uit hun loopbaananalyses blijkt dat de aanvullende verklaringswaarde van intelligentie, bovenop die van de prestaties, slechts beperkt is. In het latere cohort VOCL'99 is een op enkele onderdelen aangepaste versie van de GIVO afgenomen; deze meet meer de

‘crystalized intelligence’, dat wil zeggen de meer onder invloed van leerervaringen uitgekristalliseerde intelligentie. Deze test is speciaal ontwikkeld om schoolsucces te voorspellen. De GIVO krijgt op één na (het onderdeel ‘normen’) op alle COTAN-criteria het predicaat ‘goed’. De GIVO bestrijkt veel meer domeinen dan de PSB, maar vergt daarom ook 2 uur afnametijd. Om die reden is besloten de test in leerjaar 2 af te nemen. In de publicaties over VOCL’99 is vooralsnog geen informatie opgenomen over de resultaten van de GIVO. Wel wordt door Van der Werf, Kuyper & Zijssling (in Bosker & Meijnen, 2004) opgemerkt dat, hoewel de GIVO een verbetering is ten opzichte van de PSB, het de vraag is of het nodig is alle zeven onderdelen af te nemen. Er zou ook met drie delen volstaan kunnen worden, namelijk één van de verbale delen, één van de nonverbale delen en het uitslagendeel. Maar wellicht zou de GIVO ook helemaal kunnen vervallen omdat hij geen betere voorspeller is dan advies en grotendeels ‘crystallized intelligence’ meet. Een alternatief zou volgens Van der Werf c.s. de DAT zijn, die – net als de PSB – meer ‘fluid intelligence’ meet. Deze test is echter zeer uitgebreid (9 subtests met een gezamenlijke afnametijd van 4 uur), waardoor het nodig zal zijn een selectie te nemen. Een nadeel is ook dat deze test bedoeld is voor 13-jarigen en ouder, niveau eerste klas mavo en hoger (Evers, Van Vliet-Mulder & Groot, 2000a). Te verwachten valt dat dit problemen (bodemeffecten) zal opleveren voor een groot deel van de huidige vmbo-leerlingen.

### **3. Alternatieven voor de PRIMA-intelligentietests**

Eerder is enige twijfel geuit over opname van de PRIMA-tests in het nieuwe cohortonderzoek. De vraag is of er eventueel alternatieven zijn. Om die vraag te beantwoorden zijn enkele inventarisaties verricht.

Binnen LEO is destijds een aantal criteria gehanteerd waaraan de intelligentietests zouden moeten voldoen (Tesser, Mulder & Van der Werf, 1991):

1. geschikt voor de te onderzoeken groepen leerlingen (groep 4, 6 en 8 basisschool);
2. klassikaal afneembaar;
3. zo objectief mogelijke toekenning scores;
4. voldoende betrouwbaar;
5. beperkte afnameduur (maximaal 1 uur);
6. evenwichtige vertegenwoordiging verbale en non-verbale onderdelen;
7. minimale bias ten opzichte van allochtone leerlingen.

Uit een destijds uitgevoerde inventarisatie bleek dat veel van de in principe in aanmerking komende tests wel aan enkele criteria voldeden, maar dat er geen aan alle criteria voldeed. Om die reden is een nieuwe test geconstrueerd, waarvoor overigens in sterke mate gebruik is gemaakt van bestaande tests (zoals de PMA, GALO en DAT).

Ten behoeve van PRIMA is opnieuw een inventarisatie verricht van mogelijke instrumenten (Van Gennip, 1993). Een verschil met LEO was dat hier de doelgroep de leerlingen in groep 2 waren (en niet de groepen 4, 6 en 8). Bij de beoordeling van geschikte instrumenten zijn naast de wetenschappelijke criteria van het NIP (Evers, Van Vliet-Mulder & Ter Laak, 1992) ook pragmatische criteria gehanteerd:

1. geschikt voor groep 2;
2. voorkeur groepsgewijze afname;

3. beperkte afnameduur (i.v.m. geringe aandachtsspanne jonge kinderen);
4. kosten materiaal.

Uit deze inventarisatie bleek dat geen van de tests voldeed. De UKKI kon weliswaar groepsgewijs worden afgenomen, maar deze test voldeed niet aan alle NIP-beoordelingscriteria. De RAKIT kwam wel goed uit de bus, maar was niet groepsgewijs afneembaar. Alle andere tests waren evenmin groepsgewijs afneembaar.

Voor het nieuw op te zetten cohort is opnieuw een inventarisatie verricht. Hiervoor is gebruik gemaakt van de nieuwste versie van de COTAN-documentatie (Evers, Vliet-Mulder & Groot, 2000a,b). De gehanteerde criteria waren dezelfde als de hierboven genoemde. Ook hieruit kwam geen geschikte test naar voren. Allereerst bleek dat er nauwelijks groepsgewijs af te nemen tests beschikbaar zijn, en voorzover die er wel zijn, voldeden ze niet aan de wetenschappelijke criteria en dienden ze bovendien afgenomen te worden door een psycholoog of (ortho)pedagoog.

Onlangs is de NSCCT, de Niet-Schoolse Cognitieve Capaciteiten Test (Van Batenburg & Van der Werf, 2004; 2005) op de markt gekomen. Dit is in feite de LEO-test in een nieuw jasje. Zoals eerder opgemerkt zijn de twee in PRIMA gebruikte intelligentietests onderdelen van die LEO-test. De NSCCT is door de COTAN op alle onderdelen als ten minste 'voldoende' beoordeeld. De auteurs zelf zijn echter vrij kritisch over de kwaliteit van de test.<sup>8</sup> Volgens hen is de betrouwbaarheid van de (afzonderlijke) subtests onvoldoende, wordt de veronderstelde factorstructuur niet teruggevonden, zijn de intelligentiescores weinig stabiel, zijn de verbale tests biased ten aanzien van allochtone leerlingen, en voldoet de ruimtelijke en de numerieke factor niet aan de eisen waaraan een intelligentietest moet voldoen. Een positief punt is wel dat de score op de test een redelijk goede voorspeller van de latere prestaties en het advies voortgezet onderwijs is (predictieve validiteit). Al met al menen de auteurs dat de kwaliteit van de NSCCT wel voldoende is om de leerkrachten te helpen bij het objectief inschatten van de mogelijkheden van hun leerlingen, maar dat die onvoldoende is om ingrijpende, individuele beslissingen te nemen.

#### **4. De test van groep 4 of groep 6 voor groep 5?**

Uit het voorafgaande blijkt dat er wat betreft het basisonderwijs eigenlijk geen geschikte alternatieven zijn voor de PRIMA-test. Als toch een intelligentietest wordt opgenomen in de nieuwe cohortstudie, kan wellicht toch het beste worden teruggegrepen op de PRIMA-test. De vraag is vervolgens wel in welke jaargroepen welke test moet worden afgenomen. In de nieuwe cohort worden drie de toetsen groepen onderscheiden, namelijk groep 2, 5 en 8 (in PRIMA warend at groep 2, 4, 6 en 8). Uit de uitgevoerde inventarisaties is gebleken dat er geen geschikte test voorhanden is voor groep 2. Voor groep 8 zou de bestaande PRIMA-test kunnen worden gebruikt. De vraag is welke test in groep 5 afgenomen zou moeten worden, die van groep 4 of 6? In het ene geval bestaat het gevaar van plafondeffecten en in het andere van bodemeffecten. Om

---

<sup>8</sup> Dat de auteurs zo kritisch zijn heeft waarschijnlijk te maken met het perspectief van waaruit ze de test hebben beoordeeld, namelijk dat van individuele leerlingen. In het geval dat een dergelijke test in onderzoek wordt gebruikt voor groepsgewijze toepassingen en hij geen individuele gevolgen heeft, worden doorgaans minder strenge normen gehanteerd.

daar meer zicht op te krijgen worden in Tabel B1 en B2 de *p*-waarden en aantallen goed gemaakte items weergegeven.

*Tabel B1 – P-waarden en standaarddeviaties intelligentietest totaal groep 4 en groep 6, naar etniciteit (PRIMA-5, representatieve steekproef<sup>9</sup>)*

	Groep 4		Groep 6	
	<i>P</i>	<i>SD</i>	Gem	<i>SD</i>
Ned	.77	.14	.77	.13
Gemengd	.76	.14	.77	.13
Sur/Ant	.68	.14	.68	.17
Tur	.69	.15	.71	.15
Mar	.66	.16	.69	.15
Overig	.71	.16	.74	.14
Totaal	.75	.14	.76	.14

*Tabel B2 – Gemiddeld aantal goed gemaakte items en standaarddeviaties intelligentietest totaal groep 4 en groep 6, naar etniciteit (PRIMA-5, representatieve steekproef)*

	Groep 4		Groep 6	
	Gem	<i>SD</i>	Gem	<i>SD</i>
Ned	28.3	5.1	26.2	4.4
Gemengd	28.1	5.1	26.3	4.5
Sur/Ant	25.2	5.1	23.1	5.6
Tur	25.4	5.7	24.1	5.0
Mar	24.4	6.0	23.3	5.2
Overig	26.4	5.9	25.1	4.9
Totaal	27.9	5.3	25.9	4.6

Wat betreft de *p*-waarden wordt doorgaans voor vierkeuze-items een optimale waarde van .625 aangehouden (Dousma & Horsten, 1989; zie ook Veldhuijzen, Goldebeld & Sanders, 1993). Tabel B1 laat zien dat die waarde voor zowel groep 4 als groep 6 (lichtelijk) wordt overschreden. Dat geldt voor elk van de onderscheiden etnische groepen, maar het meest voor de Nederlandse en gemengd autochtone-allochtone groep.

In het proefonderzoek met deze tests is destijds gecontroleerd op mogelijke plafond- en bodemeffecten (vgl. Driessen, Van Langen & Oudenhoven). Daarbij is een uitsplitsing gemaakt naar 1.00-, 1.25- en 1.90-leerlingen (resp. niet-achterstandsleerlingen, autochtone en allochtone achterstandsleerlingen). De conclusie was dat er lichte plafondeffecten optraden voor de 1.00-leerlingen. Voor de meeste recente PRIMA-data (PRIMA-5) is deze exercitie herhaald, maar nu op basis van de etniciteit van de leerlingen. Daarvoor is de door het Cito gehanteerde formule gehanteerd: geen plafondeffect als  $k-(\text{gem}+1.5*sd)>.1*k$ , en geen bodemeffect als  $k-(\text{gem}-1.5*sd)>.1*k$ , waarbij *k* het aantal items representeert. Toegepast op Tabel B2 laten de resulta-

<sup>9</sup> De parameters voor de totale steekproef wijken hooguit een fractie af van die van de representatieve steekproef.



ten zien dat er lichte plafondeffecten optreden voor de Nederlandse en gemengd autochtone-allochtone leerlingen in groep 4 en 6; er zijn geen bodemeffecten.

Deze bevindingen pleiten er voor de test van groep 6 te nemen om in groep 5 te worden afgenomen. Bosker & Meijnen (2004) stellen dat er voor PRIMA een andere test nodig is vanwege de daar gesignaleerde plafondeffecten. Door voor groep 5 de toets van groep 6 te nemen wordt voor die groep dat probleem opgelost.

## **5. Nabeschuwing**

De vraag is of er een intelligentietest moet worden opgenomen in de nieuwe cohortstudie, en zo ja, welke en wanneer. Tot nu toe zijn in het primair onderwijs onderdelen van de LEO-test afgenomen. Technisch gezien voldoet deze test redelijk. Echter, als effectvariabele heeft hij nauwelijks meerwaarde ten opzichte van prestatietoetsen, en als controlevariabele evenmin. In het voortgezet onderwijs is eerst de PSB afgenomen. Deze test meet evenals de LEO-test vooral de ‘fluid intelligence’ ofwel zoiets als aanleg. Ook voor deze test geldt dat hij nauwelijks meerwaarde heeft. Bovendien zijn er twijfels met betrekking tot de inhoudsvaliditeit. Om die reden is deze test vervangen door (onderdelen van) de GIVO. Deze test is veel meer op ‘crystalized intelligence’ gericht, ofwel het resultaat van scholing. Echter, nu wordt door de betreffende onderzoekers geconcludeerd dat ook deze test niet voldoet en wordt ter overweging gegeven weer over te stappen op (c.q. terug te keren naar) een andere test die ‘fluid intelligence’ meet. De test die wordt aanbevolen (de DAT) lijkt echter niet helemaal geschikt voor de doelgroep.

Bij dit alles is onduidelijk wat het doel van de test zou moeten zijn, dient hij de begincapaciteiten/aanleg te meten of (veel meer) het resultaat van het onderwijs? Ofwel, moet hij als controle- of als effectvariabele worden beschouwd? Als hij als controlevariabele moet worden opgevat, dan ligt het voor de hand de test zo vroeg mogelijk af te nemen en volstaat een eenmalige afname. Als hij als effectvariabele moet worden gezien, dan zou afname op een later moment voor de hand liggen. Beide functies zouden ook gecombineerd kunnen worden, maar dat zou de inzet van twee verschillende tests vereisen.

Kortom, gegeven de ervaringen met de tests in PRIMA en VOCL, de zeer beperkte meerwaarde, zowel in termen van controle- als effectvariabele, het niet-beschikbaar zijn van geschikte alternatieven en de onduidelijkheid wat betreft de functie is het de vraag of er überhaupt een intelligentietest zou moeten worden meegenomen. Als er al een zou moeten worden afgenomen, dan zou dat waarschijnlijk een test moeten zijn die ‘fluid intelligence’ meet, die verschillende domeinen bestrijkt en die in de beginfase van het primair onderwijs kan worden afgenomen. Aangezien een dergelijke test voor gebruik in het beoogde cohortonderzoek niet beschikbaar is, is het misschien beter helemaal geen intelligentietest mee te nemen.

